



Modeling and Simulation: The Good, the Bad, and the Hopeful

David B. Nelson, Ph.D.

Director

National Coordination Office for
Information Technology Research and Development

*DOE Computational Science Graduate Fellowship
Conference*

July 15, 2003



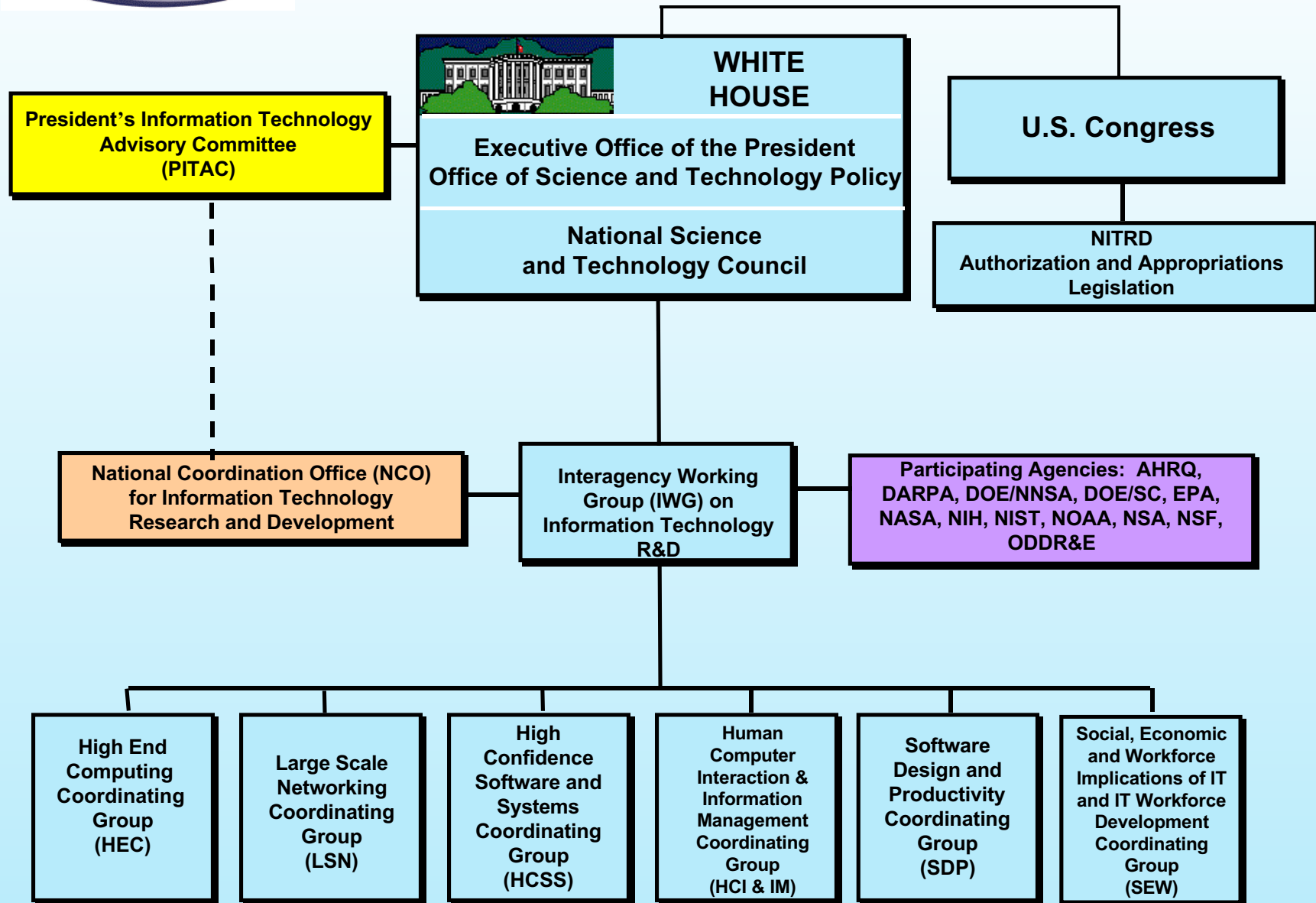
National Coordination Office (NCO) for Information Technology Research and Development (IT R&D)

Mission: *To formulate and promote Federal information technology research and development to meet national goals.*

- NCO Director reports to the Director of the White House Office of Science Technology Policy (OSTP) and co-chairs the Interagency Working Group for IT R&D
- Coordinates planning, budget, and assessment activities for the Federal multi-agency Networking and Information Technology R&D (NITRD) Program
- Supports six technical Coordinating Groups (CGs) that report to the Interagency Working Group



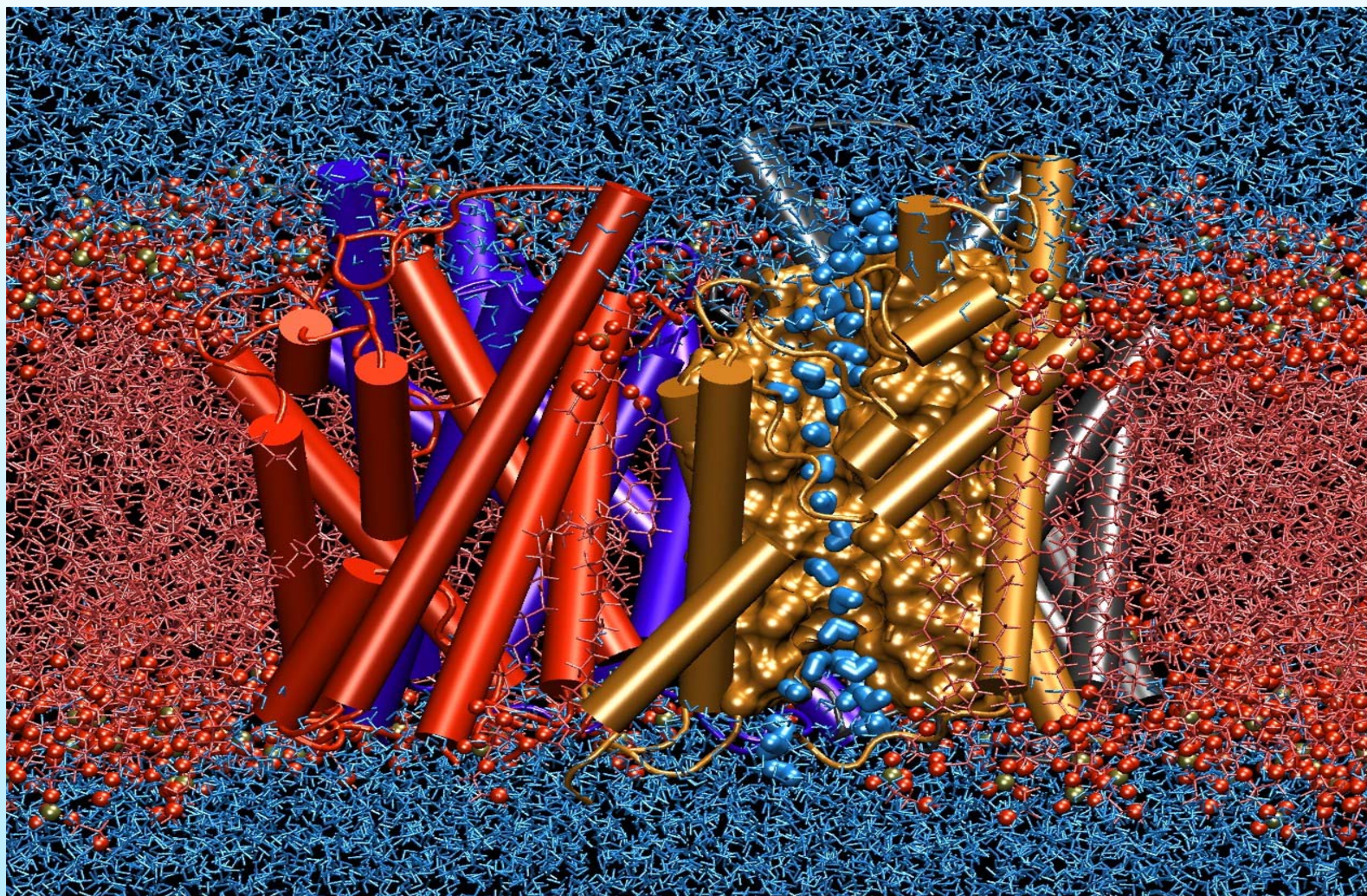
NITRD Program Coordination





Simulation of Aquaporin Protein Inside a Cell (NSF, NIH, PSC Alpha Cluster)

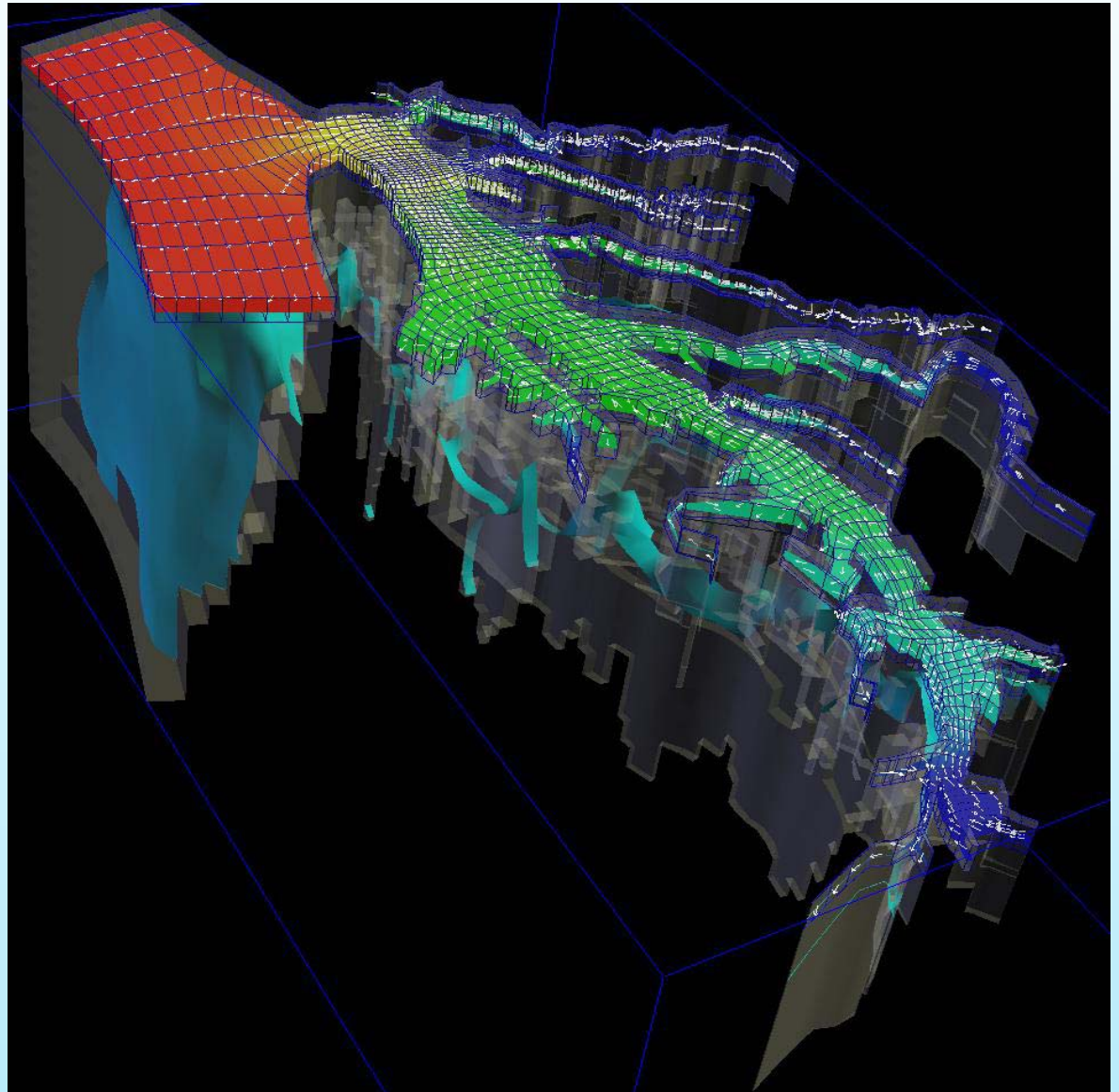
Visualization shows transport of water molecules into cell.





Environmental Modeling of the Chesapeake Bay (NOAA, EPA, DoD)

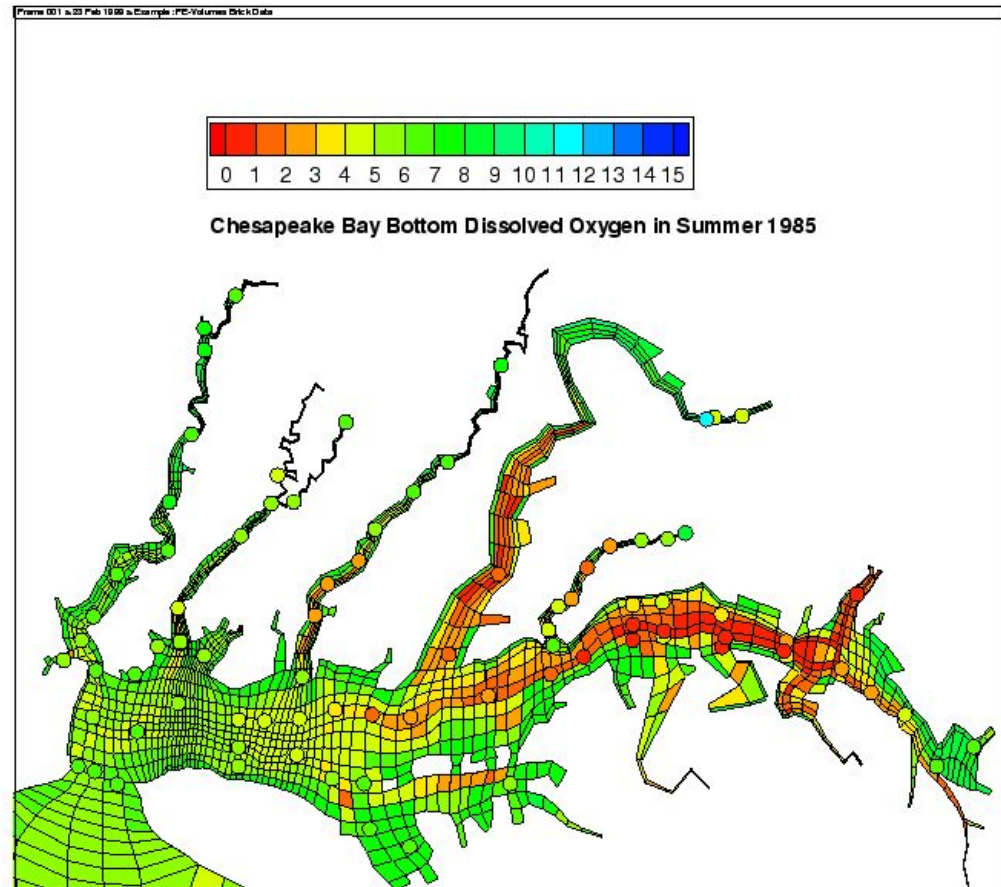
- Image shows visualization of computed salinity in the Bay (red is high salinity.)
- South is up.
- Visualization is an important part of the model, because users may not be skilled computational scientists.





Environmental Modeling of the Chesapeake Bay (NOAA, EPA, DoD)

- Model is checked against measured data
- Model has shown that approximately 1/4 of the nitrogen added to the Bay starts as air pollution, some from sources hundreds of miles from the Bay's watershed.
- Model also shows that substantial nitrogen comes from ground water on the Eastern shore

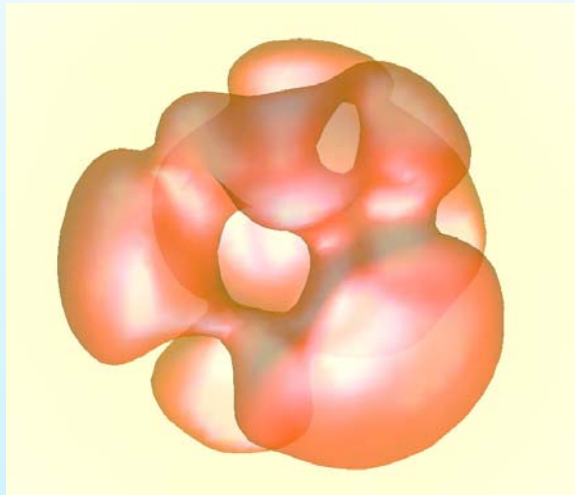




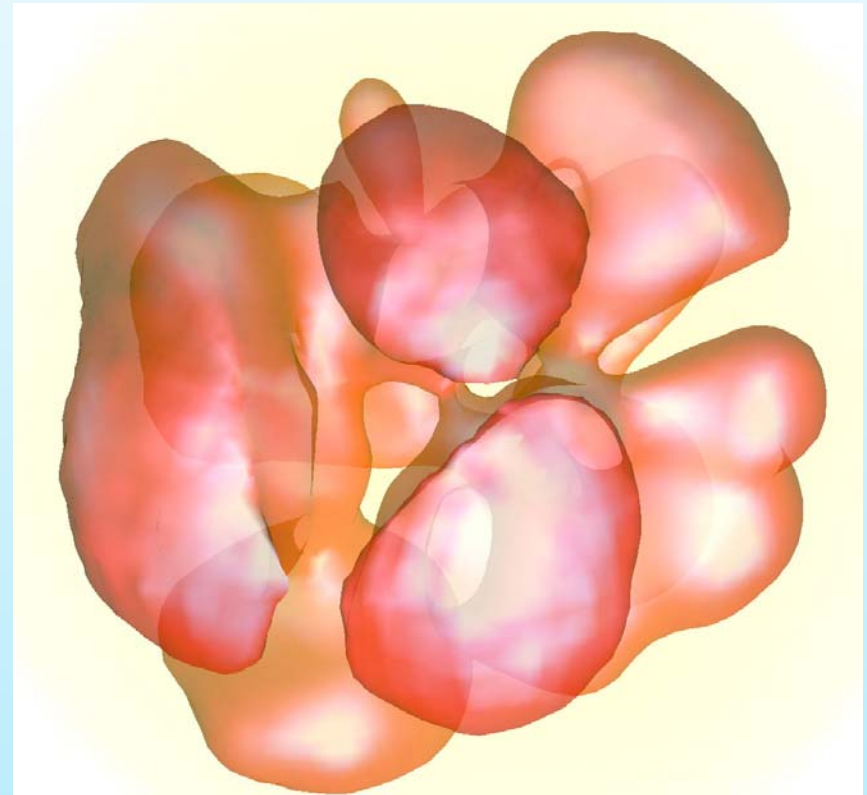
Explosion of a Super-Nova (not to scale) (DOE)



Start

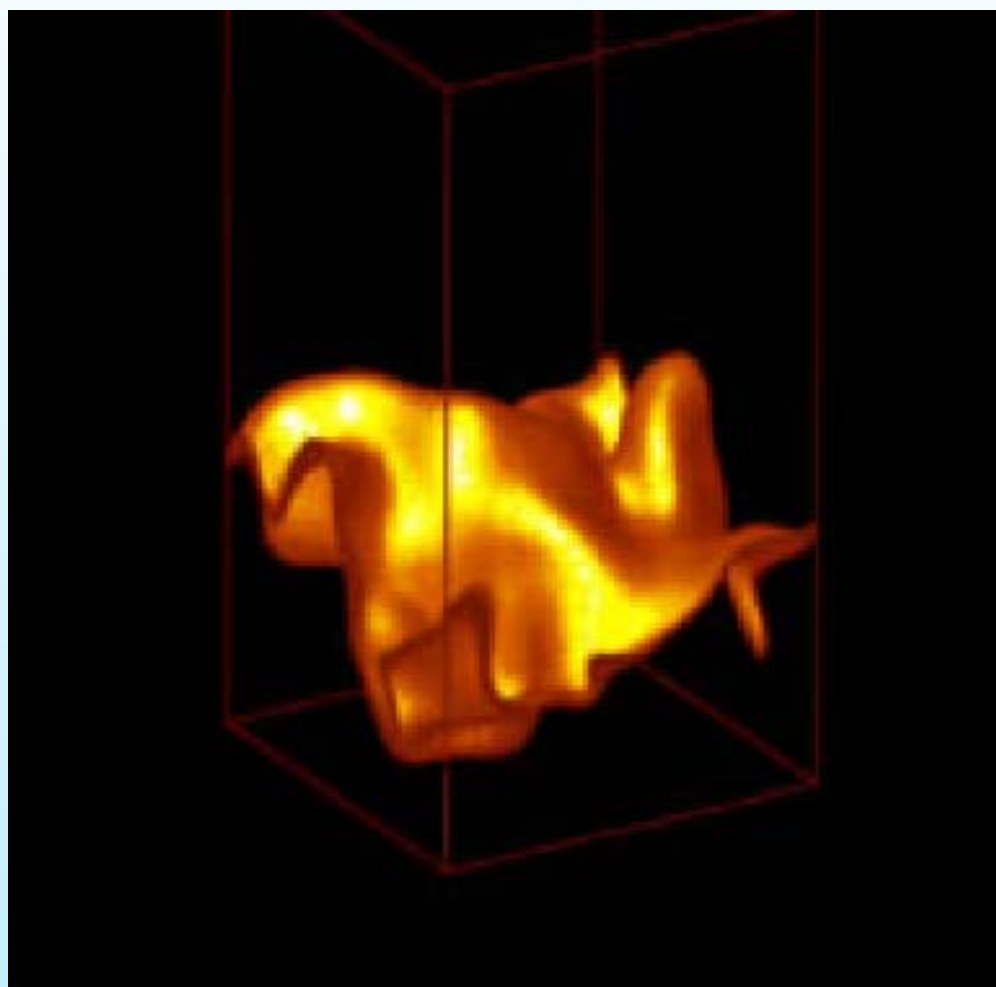


Middle



End

Simulation of Turbulent Flame with Comprehensive Chemistry (DOE)

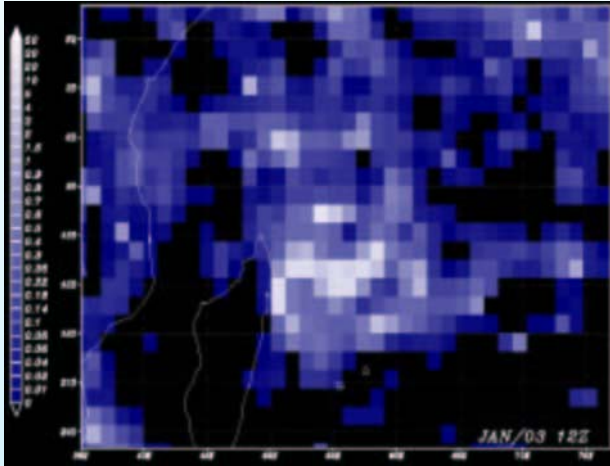


Flame surface from simulation of a turbulent premixed methane flame

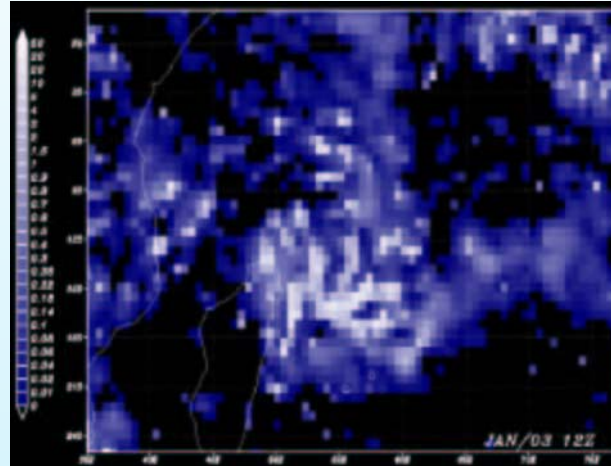


Power of Japanese Earth Simulator Allows Better Resolution of Local Features

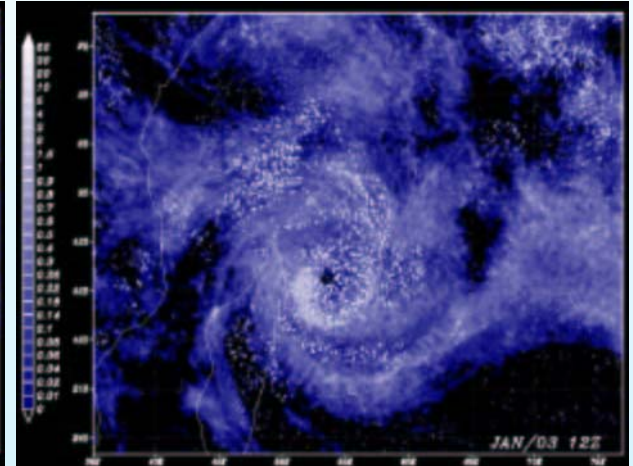
Simulation of Tropical Cyclone Near Madagascar



125.1 km grid



62.5 km grid



10.4 km grid

(U.S. 1200 year control run used approximately 280 km grid.)



Grid Communications & Applications: High End Physics Problem



Compact Muon Solenoid at CERN

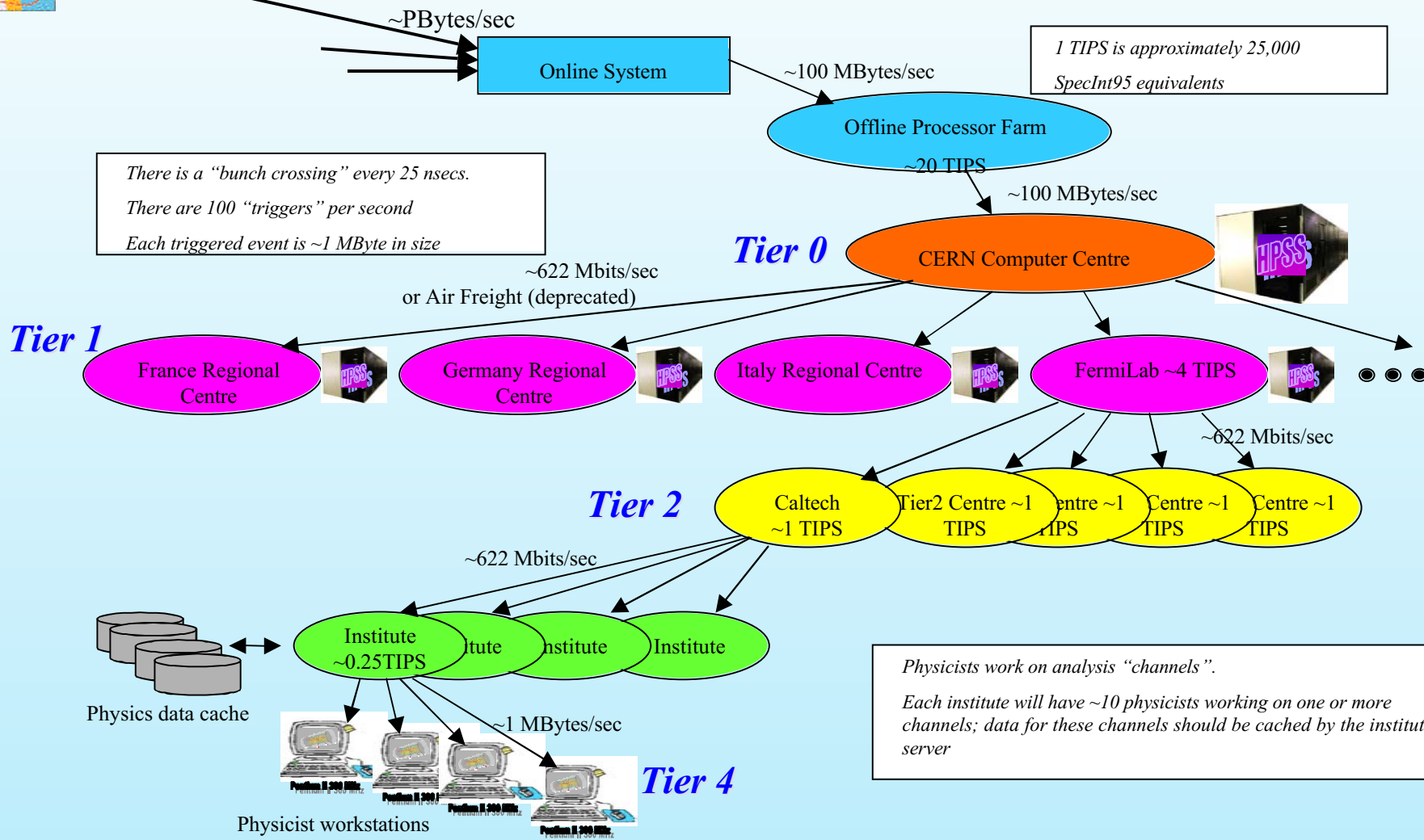
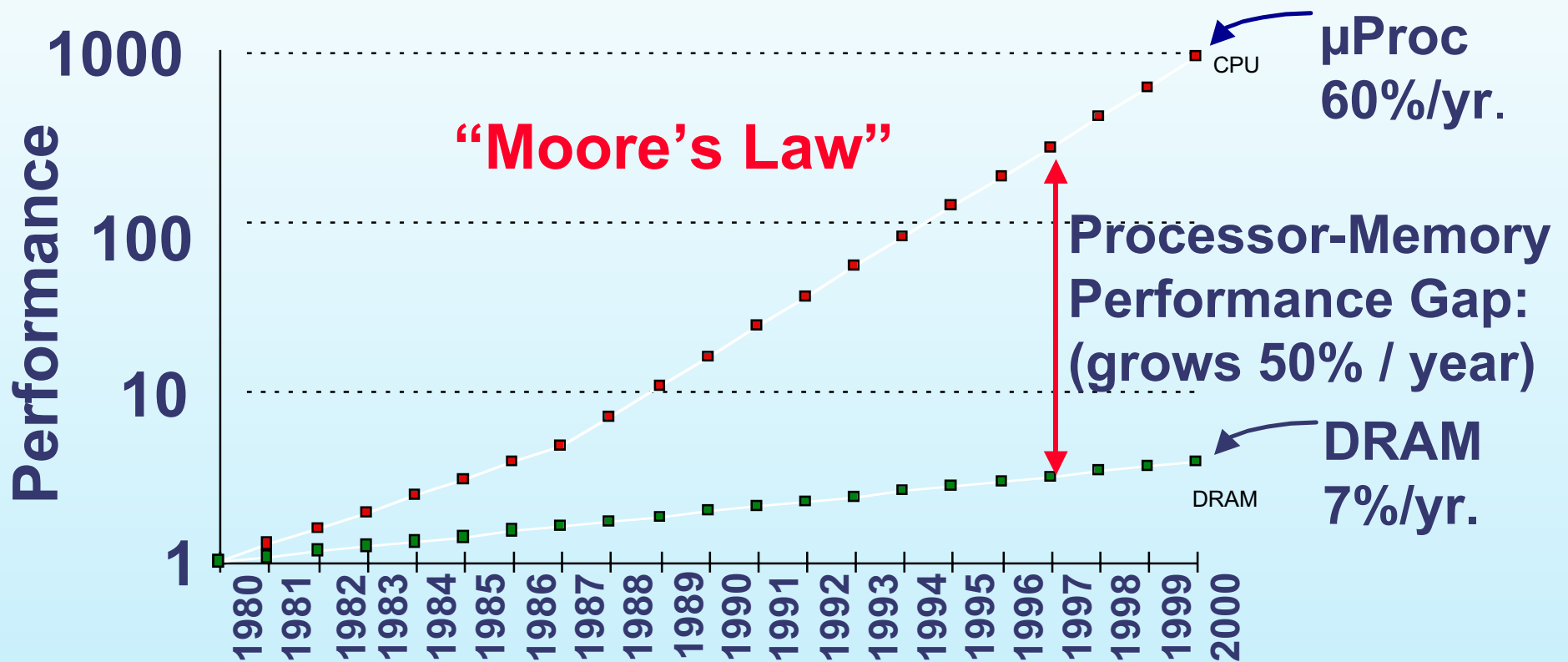


Image courtesy Harvey Newman, Caltech

Processor-Memory Performance Gap



- Alpha 21264 full cache miss / instructions executed:
180 ns/1.7 ns = 108 clks x 4 or 432 instructions
- Caches in Pentium Pro: 64% area, 88% transistors

*Taken from Patterson-Keeton Talk to SigMod



Processing vs. Memory Access

- **Doesn't cache solve this problem?**

It depends. With small amounts of contiguous data, usually. With large amounts of non-contiguous data, usually not.

In most computers the programmer has no control over cache.

Often “a few” Bytes/FLOP is considered OK.

- **However, consider operations on the transpose of a matrix (e.g., for adjoint problems)**

$$Xa = b$$

$$X^T a = b$$

If X is big enough, 100% cache misses are guaranteed, and we need at least 8 Bytes/FLOP (assuming a and b can be held in cache).

- **Latency and limited bandwidth of processor-memory and node-node communications are major limiters of performance for scientific computation**



Testing Processing vs. Memory Access with Benchmarks

- **Simple benchmark: Stream Triad**

$$a_i + s \times b_i = c_i$$

a_i , b_i , and c_i are vectors; s is a scalar. Vector length is chosen to be much longer than cache size.

Each execution includes 2 memory loads + 1 memory store and 2 FLOPs, or 12 Bytes/FLOP (assuming 8 Byte precision)

No computer has enough memory bandwidth to reference 12 Bytes for each FLOP!



Testing Processing vs. Memory Access with Benchmarks

- **Another Benchmark: Linpack**

$$A_{ij} x_j = b_i$$

Solve this linear equation for the vector x , where A is a known matrix, and b is a known vector. Linpack uses the BLAS routines, which divide A into blocks.

**On the average Linpack requires 1 memory reference for every 2 FLOPs, or 4Bytes/Flop.
Many of these can be cache references.**



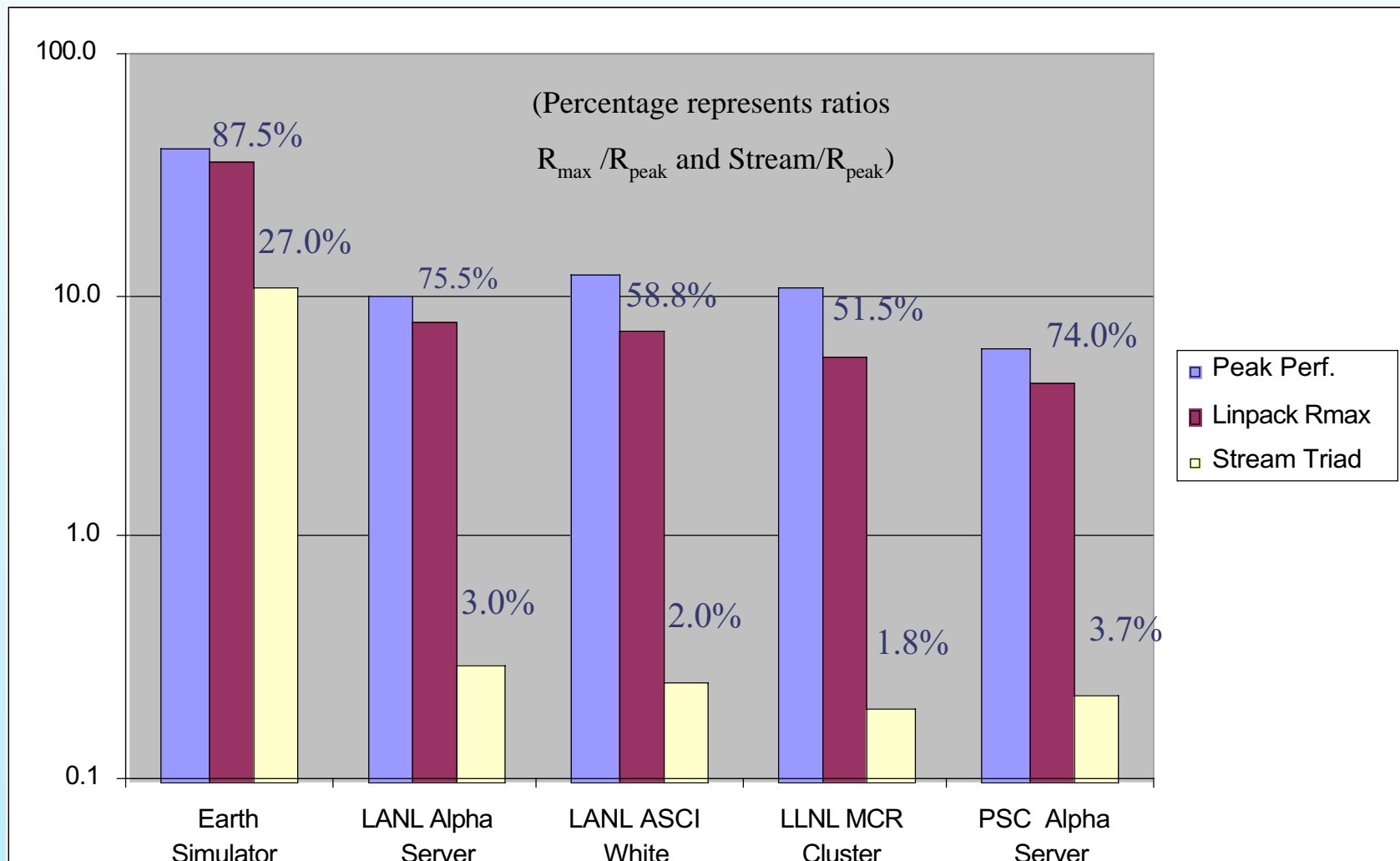
Selected System Characteristics

| | | Earth Simulator (NEC) | ASCI Q (HP ES45) | ASCI White (IBM SP3) | MCR (Dual Xeon) | Cray X1 (Cray) |
|---------------------------|----------------------|------------------------------|----------------------------|-------------------------|----------------------------|--------------------------|
| Year of Introduction | | 2002 | 2003 | 2000 | 2002 | 2003 |
| Node Architecture | | Vector SMP | Alpha micro SMP | Power 3 micro SMP | Xeon micro SMP | Vector SMP |
| System Topology | | NEC single-stage Crossbar | Quadrics QsNet Fat-tree | IBM Omega network | Quadrics QsNet Fat-tree | 2D Torus Interconnect |
| Number of Nodes | | 640 | 3072 (Total) | 512 | 1152 | |
| Processors | - per node | 8 | 4 | 16 | 2 | 4 |
| | - system total | 5120 | 12288 | 8192 | 2304 | |
| Processor Speed | | 500 MHz | 1.25 GHz | 375 MHz | 2.4 GHz | 800 MHz |
| Peak Speed | - per processor | 8 Gflops | 2.5 Gflops | 1.5 Gflops | 4.8 Gflops | 12.8 Gflops |
| | - per node | 64 Gflops | 10 Gflops | 24 Gflops | 9.6 Gflops | 51.2 Gflops |
| | - system total | 40 Gflops | 30 Gflops | 12 Tflops | 10.8 Tflops | |
| Memory | - per node | 16 GB | 16 GB | 16 GB | 16 GB | 8-64 GB |
| | - per processor | 2 GB | 4 GB | 1 GB | 2 GB | 2-16 GB |
| | - system total | 10.24 TB | 48 TB | 8 TB | 4.6 TB | |
| Memory Bandwidth (peak) | | | | | | |
| | - L1 Cache | N/A | 20 GB/s | 5 GB/s | 20 GB/s | 76.8 GB/s |
| | - L2 Cache | N/A | 13 GB/s | 2 GB/s | 1.5 GB/s | |
| | Main (per processor) | 32 GB/s | 2 GB/s | 1 GB/s | 2 GB/s | 34.1 GB/s |
| Inter-node MPI | | | | | | |
| | - Latency | 8.6 μ sec | 5 μ sec | 18 μ sec | 4.75 μ sec | |
| | - Bandwidth | 11.8 GB/s | 300 MB/s | 500 MB/s | 315 MB/s | 12.8 GB/s |
| Bytes/flop to main memory | | 4 | 0.8 | 0.67 | 0.4 | 2.66 |
| Bytes/flop interconnect | | 1.5 | 0.12 | 0.33 | 0.07 | 1 |

Most of this data is from Kerbyson, Hoisie, Wasserman; LANL; unpublished



Performance Measures of Selected Top Computers



Note Logarithmic Y axis



Major Problem: Poor Links Between Workload and Architecture Design

- **Build It and They Will Come**
- **Weakness of Government High Performance Computing and Communication Program in 1990s**
 - No link between grants for computer architecture research and grants for computer acquisition
 - Poor feedback from users to developers
 - Poor connections between computational scientists and computer scientists (one workshop in Pittsburgh in 1993)
- **Result: Selection of computer architecture is not well grounded on application needs**



What About Synthetic Benchmarks?

- Peak performance – nuf said
- Linpack –only measures performance of cache-friendly code
- Stream – only measures contiguous communications with memory, but good measure of bandwidth
- GUPS – really tough benchmark because it makes random memory access; may exceed requirements of most codes
- IDC balanced benchmarks – good compilation, but somewhat artificially combined
- Effective System Performance Benchmark – promising, but not widely used
- NAS Parallel Benchmarks – disused, but may be coming back
- Livermore Loops – obsolete
- Your own workload - ??



Resurgence of Performance Analysis Is Promising

- LANL Performance and Architecture Lab:
http://www.c3.lanl.gov/par_arch/
- Performance Evaluation Research Center:
<http://perc.nerisc.gov/>
- IDC User Forum: <http://64.122.81.35/benchmark/>
- Performance Modeling and Characterization:
<http://www.sdsc.edu/PMaC/Benchmark/>
- NAS Parallel Benchmarks:
<http://www.nas.nasa.gov/Software/NPB/>
- Recent High End Computing Workshop offered recommendations for performance evaluation:
<http://www.cra.orgActivities/workshops/nitrd/>
- Great opportunity for agencies to cooperate on performance evaluation.



Summary

- **Computational Science is now a third pillar of research, along with experiment and theory.**
- **High-end computers are getting harder to use and more inefficient.**
- **Federal agencies are recognizing this and working to improve things.**



For Further Information

Please contact us at:

nco@itrd.gov

Or visit us on the Web:

www.itrd.gov